TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 23, No. 1, February 2025, pp. 129~137
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v23i1.25341

Prediction of heart disease using random forest algorithm, support vector machine, and neural network

Didik Setiyadi[1], Henderi[2], Anrie Suryaningrat[3], Rulin Swastika[4], Saludin[5], Muhamad Malik Mutoffar[6], Imam Yunianto[7]

[1]Department of Information System, STMIK Sinar Nusantara, Surakarta, Indonesia
[2]Department of Informatics Engineering, Faculty of Science and Technology, Universitas Raharja, Tangerang, Indonesia
[3]Department of Informatics Engineering, Faculty of Engineering and Informatics, Universitas Dian Nusantara, Jakarta, Indonesia
[4]Department of Informatics Engineering, Faculty of Computer Science, Universitas Al-Khairiyah, Banten, Indonesia
[5]Department of Informatics Engineering, Faculty of Informatics, Universitas Bina Insani, Bekasi, Indonesia
[6]Department of Informatics Engineering, Faculty of Creative Industries, Universitas Teknologi Bandung, Bandung, Indonesia
[7]Department of Informatics Engineering, Faculty of Engineering and Communication, Institut Bisnis Muhammadiyah Bekasi, Bekasi, Indonesia

Article Info

Article history:

Received May 17, 2024
Revised Aug 3, 2024
Accepted Sep 20, 2024

Keywords:

Heart disease
Neural network
Prediction
Random forest
Support vector machine

ABSTRACT

The heart is a vital organ responsible for pumping blood throughout the human body. Machine learning has become an increasingly important tool in medical forecasting, improving diagnostic accuracy and reducing human errors. This study focuses on detecting heart disease using machine learning algorithms. It aims to compare the performance of three key algorithms random forest (RF), support vector machine (SVM), and neural networks (NN), in predicting heart disease. Using a patient dataset with both nominal and numeric attributes, record mining techniques were applied through Orange software. The target classes indicated the absence (0) or presence (1) of heart disorders. The evaluation was based on the prediction accuracy of each algorithm. Results show that SVM achieved the highest accuracy, with a rate of 85%, outperforming RF and NN. The findings suggest that the SVM algorithm is a reliable tool for heart disease prediction, helping reduce diagnostic errors and improve medical decision-making.

Corresponding Author:

Muhamad Malik Mutoffar
Department of Informatics Engineering, Faculty of Creative Industries
Universitas Teknologi Bandung
Bandung, Indonesia
Email: malik@utb-univ.ac.id

1. INTRODUCTION

Smart health system as "the result of a natural synergy between m-health and smart cities, from an information and communication technology (ICT) perspective as well as individuals and communities", m-Health stands for healthcare services via mobile devices [1]. Several implementations of e-Healt and m-Health are being developed continuously to implement a prediction system and find the best algorithm for several diseases, especially heart disease. The coronary heart is a essential organ that is owned by humans, and it pumps blood throughout the body. If the heart is problematic then humans will get many diseases that attack the body, therefore humans must have a healthy heart.

Under relaxed conditions, the normal heart rate of humans over the age of 18 is generally between 60 beats per minute (BPM) to 100 BPM, pumping about 2,000 gallons of blood every day and it is the most

important organ in the human body. Chronic heart disease and disease progression occurs when the heart cannot supply sufficient blood to tissues and organs at physiological filling pressures to meet metabolic needs. Despite advances in medicine and the healthcare industry, the five-year survival rate for people affected by heart disease is around 50% [2], [3]. The development of heart disease is one causes disability disruption of balance between blood supply and demand that occurs due to blockage of blood vessels [4], [5]. This category of heart disease is called circulatory disease. Various types of heart disease including congenital heart disease, dilated cardiomyopathy, heart failure, coronary artery disease, myocardial infarction, mitral regurgitation, hypertrophic cardiomyopathy, arrhythmia [6]. The death rate from heart disease is among the highest in the world.

Lack of skilled doctors and ignoring the patient's symptoms causes great challenges that can lead to death, not good for patients [7], [8]. Therefore, an information system is needed analysis tool to find hidden patterns in heart disorder data. Data minning is a cognitive process to find hidden proximal patterns in large amounts of data [9], [10]. Here, various techniques are listed and tested to determine the accuracy level of each. In previous studies, researchers have expressed efforts to find the best predictive model. The use of machine learning is rapidly increasing in the medical diagnostics industry as computer analysis reduces manual errors and improves accuracy [11]. Using machine learning techniques, disease diagnosis becomes highly reliable [12].

Medical data mining is a very important research area, it has contributed to the development of various applications in the growing healthcare field [13], [14]. Data mining is the search for useful information from very large sets of data. Some of the most important and common techniques for data mining include association rules, classification, clustering, prediction, and sequential models [15]. Data mining techniques are used for various applications. Data mining plays an important role in disease detection in healthcare. Patients should be asked to undergo a series of tests for the disease, although data mining techniques may be used to reduce the number of tests.

It has been stated in the literature that applying technical features and feature selection improves both classification and prediction results. Various machine learning and deep learning techniques are used to detect cardiac disease and hyperparameter optimization is performed to increase the accuracy of the results. Neural networks (NN) achieve a high accuracy of 78.3 percent and other models include ensemble techniques such as logistic regression, support vector machine (SVM), and random forest (RF) to mitigate cardiovascular features [16]. Various data mining techniques can be used to analyze cardiac related problems. There are various types of data mining techniques such as decision trees (DT), Naïve Bayes (NB), SVM, K-nearest neighbor (K-NN) classifiers, hybrid approaches, and artificial neural networks (ANN). Classification algorithms such as NB, DT, and ANN have been widely used in predicting heart diseases and have achieved different levels of accuracy [17], [18].

This study compares the results of the best algorithms from several previous studies, RF algorithm, SVM, and ANN, to get the best algorithm for predicting heart disease. The main objective of this study was to identify the best classification algorithm from previous studies conducted by [12]. For this reason, data mining techniques are carried out by comparing three data mining classification algorithms, namely the RF algorithm, SVM, and NN.

2. METHOD

This research involves an investigation of the treatment of parameters and variables which all depend on the researcher himself. Heart disease dataset from the latest University of California, Invene (UCI) repository on September 11, 2021. UCI machine learning consists of four databases extracting information from Hungary, Cleveland, Long Beach, Virginia, and Switzerland. A total of 918 datasets with a total of 76 attributes are included, but a subset of 12 attributes was used in the experiments. These attributes are categorical, real-valued, and integer-valued. The dataset attributes can be seen as in Table 1.

The collection of research literature from articles or journals adapted to this research is the prediction of heart disease. For the next step is choosing models or algorithms from previous studies, namely the RF algorithm, SVM and NN. Furthermore, training for the dataset is given, then prediction evaluation with area under the curve (AUC), component analysis (CA), F1, precision, recall, confusion matrix and receiver operating characteristic (ROC) analysis. The process can be seen in Figure 1. The data mining process written in Figure 1 is implemented using the Orange application as shown in Figure 2.

Table 1. Attributes of the heart disease dataset [11]

Attribute	Information attribute	Description
Age	Integer	Age in years (29 to 77)
Sex	Integer	Gender instance (0=Female, 1=Male)
ChestPainType	Integer	Chest pain type (1: typical angina, 2: atypical angina, 3: non-anginal pain, and 4:asymptomatic)
RestingBP	Integer	Resting blood pressure in mm Hg
Cholesterol	Integer	Serum cholesterol in mg/dl
FastingBS	Integer	Fasting blood sugar >120 mg/dl (0=False, 1=True)
RestingECG	Integer	Resting ECG results (0: normal, 1: ST-Twave abnormality, 2: LV hypertrophy)
MaxHR	Integer	Maximum heart rate achieved
ExerciseAngina	Integer	Exercise induced angina (0: No, 1: Yes)
Oldpeak	Integer	ST depression induced by exercise relativeto rest [0.0, 62.0]
ST_Slope	Integer	Slope of the peak exe+A1:D13rcise ST segment (1: up-sloping, 2: flat, 3: down-sloping)
HeartDisease	Integer	Target

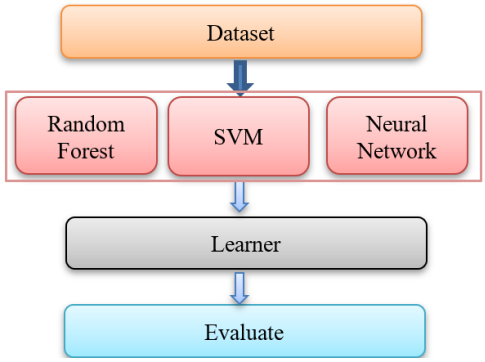


Figure 1. Data mining process

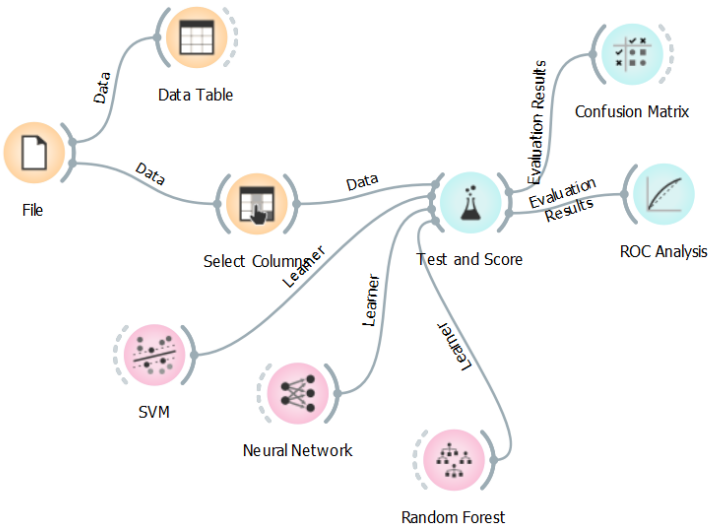


Figure 2. Workflow of heart disease prediction process

2.1. Random forest

The RF algorithm is a collection of classifiers for various tree types. Leo Breiman pioneered the use of randomized forests in formal studies. He explains how to generating a forest of unrelated trees in his work by combining the classification and regression treelike (CART-like) procedure with random node optimization and bagging. The Breiman algorithm aims to combine decisions from several rather than generating a single DT, multivariate trees, each trained on a different training cluster [19].

The RF algorithm uses several DT to determine the classification value during the classification process. Instead of dividing each node into branches based on the best branch among all variables, the RF algorithm divides each node into variables based on the best variable among the variables chosen at random in each node [20]. Each dataset is generated using data that has been moved from the original. Then, using the random selection feature, the tree is developed without pruning [21], [22].

2.2. Neural network

NNs are an attempt to mimic the functioning of the human brain. The human brain is thought to be made up of millions of small processing units called neurons that operate in parallel. The neurons are connected to each other by neuronal connections. Each neuron receives input from groups of these neurons, processes the input, and passes an output back to the group of neurons [23]. The output is collected by other neurons for further processing [24]. The quantity of enter neurons withinside the enter layer is 25 neurons, one hidden layer has 50 hidden neurons, and the output layer contains one neuron to determine the hypertension class (1) or not hypertension (0) with 1,000 epochs or iterations [25]. The hidden layer used in this research is 100.

2.3. Support vector machine

SVM is a guided learning method in classifying. SVM can be used in linear or non-linear classification [26]. The best hyperplane can be found by measuring the hyperplane margins and finding the maximum point [27]-[29]. Basically non-linear SVM is a solution to the linear SVM problem by performing kernel functions in a high-dimensional feature space [30].

Some of the performance metrics that are common and often used are as follows.

a. Accuracy

Accuracy can indicate how well a model classifies correctly. Therefore, prediction accuracy is the ratio of the number of true positives and true negatives to the whole data. In other words, accuracy is how close the predicted value is to the actual value. The value of accuracy can be seen in (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b. Precision

Accuracy can be explained as the degree of accuracy of the requested data with the prediction results provided by the model. Therefore, accuracy can be interpreted as the ratio of correct positive predictions divided by the total positive prediction results. In other words, out of all the correctly predicted positive class data, how many data are actually positive? The value of accuracy can be seen in (2).

$$Presisi = \frac{TP}{TP+FP} \quad (2)$$

c. Recall

Recall or sensitivity can be described as the success of a model in retrieving information. Therefore, we can say that recall is the proportion of true positive predictions divided by the total number of true positives. The value of recall can be found in (3).

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

d. Specificity

Specificity refers to the accuracy of predicting negative data, divided by the total number of negative cases. It indicates how well a model correctly identifies negative instances. A higher specificity means the model is effective at minimizing false positives. Specificity is typically calculated in medical diagnostics to ensure that healthy individuals are not mistakenly classified as having a disease. The value of specificity can be found in (4).

$$Specificity = \frac{TN}{TN+F} \quad (4)$$

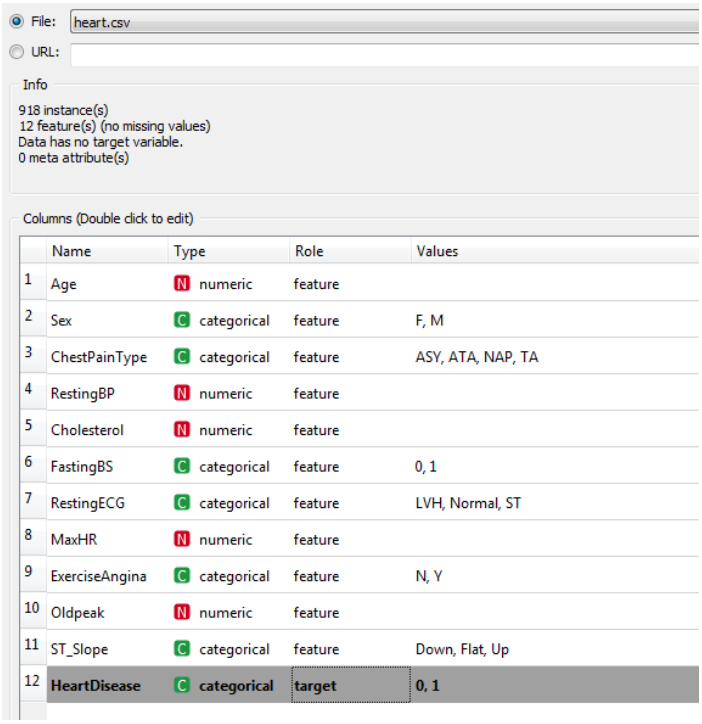
e. F1 score

F1 score is a weighted comparison of average precision and recall. F1 score is best when the system has some balance between precision and recall. The value of F1 score can be found in (5).

$$Skor\ F1 = \frac{2 \times Recall \times Presisi}{Recall + Presisi} \quad (5)$$

3. RESULTS AND DISCUSSION

The method used in this study is a classification algorithm with RF algorithm, NN and SVM, it was used to develop a predictive system to analyze and predict the likelihood of heart disease. To get the performance of this classification algorithm use publicly available datasets in the UCI repository with the cardiology dataset. The dataset used has 12 attributes consisting of nominal and numeric data. The target data classes are absence of heart disease and the presence of heart disease. Can be seen in Figure 3. In the dataset used, 11 features are selected which are used to make predictions, as shown in Figure 4.



	Name	Type	Role	Values
1	Age	numeric	feature	
2	Sex	categorical	feature	F, M
3	ChestPainType	categorical	feature	ASY, ATA, NAP, TA
4	RestingBP	numeric	feature	
5	Cholesterol	numeric	feature	
6	FastingBS	categorical	feature	0, 1
7	RestingECG	categorical	feature	LVH, Normal, ST
8	MaxHR	numeric	feature	
9	ExerciseAngina	categorical	feature	N, Y
10	Oldpeak	numeric	feature	
11	ST_Slope	categorical	feature	Down, Flat, Up
12	**HeartDisease**	**categorical**	**target**	**0, 1**

Figure 3. Dataset used

Data instances: 918
Features: 11
Meta attributes: None
Target: Class 'HeartDisease'

	HeartDisease	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope
1	0	40	M	ATA	140	289	0	Normal	172	N	0.0	Up
2	1	49	F	NAP	160	180	0	Normal	156	N	1.0	Flat
3	0	37	M	ATA	130	283	0	ST	98	N	0.0	Up
4	1	48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat
5	0	54	M	NAP	150	195	0	Normal	122	N	0.0	Up
6	0	39	M	NAP	120	339	0	Normal	170	N	0.0	Up
7	0	45	F	ATA	130	237	0	Normal	170	N	0.0	Up
8	0	54	M	ATA	110	208	0	Normal	142	N	0.0	Up
9	1	37	M	ASY	140	207	0	Normal	130	Y	1.5	Flat
10	0	48	F	ATA	120	284	0	Normal	120	N	0.0	Up
11	0	37	F	NAP	130	211	0	Normal	142	N	0.0	Up
12	1	58	M	ATA	136	164	0	ST	99	Y	2.0	Flat
13	0	39	M	ATA	120	204	0	Normal	145	N	0.0	Up
14	1	49	M	ASY	140	234	0	Normal	140	Y	1.0	Flat
15	0	42	F	NAP	115	211	0	ST	137	N	0.0	Up
16	0	54	F	ATA	120	273	0	Normal	150	N	1.5	Flat
17	1	38	M	ASY	110	196	0	Normal	166	N	0.0	Flat
18	0	43	F	ATA	120	201	0	Normal	165	N	0.0	Up
19	1	60	M	ASY	100	248	0	Normal	125	N	1.0	Flat
20	1	36	M	ATA	120	267	0	Normal	160	N	3.0	Flat
21	0	43	F	TA	100	223	0	Normal	142	N	0.0	Up
22	0	44	M	ATA	120	184	0	Normal	142	N	1.0	Flat
23	0	49	F	ATA	124	201	0	Normal	164	N	0.0	Up

Figure 4. Features used for heart prediction

This system consists of several stages, namely preprocessing to eliminate noisy or redundant data, then classifying it using the RF algorithm, NN, and SVM. After that, validation is carried out to see the performance of the algorithm, namely AUC, CA, F1, precision, and recall from each algorithm in order to get the algorithm with the best accuracy. Doing this experiment using the Orange application with a daset that has 12 thousand consisting of numeric and nominal data. The classification algorithms tested are RF algorithm, NN and SVM. The confusion matrix value calculates the accuracy of the data mining concept or decision support system, serves to analyze whether the classifier is good at recognizing different classes. The following is a confusion matrix of the RF algorithm, NN and SVM.

The RF algorithm in Figure 5 contains 3351+4511 correct predictions and 749+589 incorrect predictions from a total of 9,200 testing data. Thus, the accuracy of this algorithm (precision) can be calculated as 7862/1338*100%=59%. The SVM algorithm in Figure 6 contains 3281+4545 correct predictions and 819+555 incorrect predictions from a total of 9,200 testing data. Thus, the accuracy of this algorithm (precision) can be calculated as 7826/1374*100%=56%. The SVM algorithm in Figure 7 contains 3364+4453 correct predictions and 736+647 incorrect predictions from a total of 9,200 testing data. Thus, the accuracy of this algorithm (precision) can be calculated as 7817/1383*100%=56%.

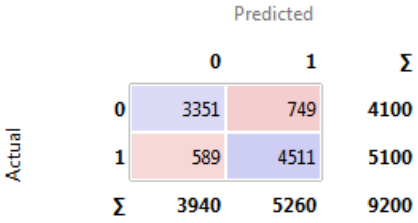


Figure 5. RF confusion matrix

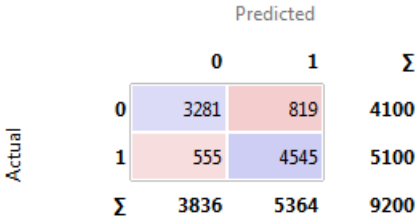


Figure 6. Confusion matrix SVM

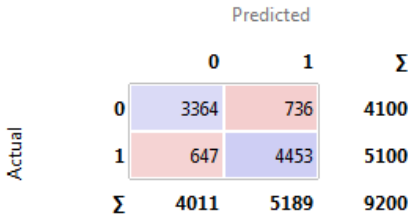


Figure 7. Confusion matrix NN

In Table 2, the value can be seen based on the formula from Table 1. AUC, CA, F1, precision, and recall are the methods used when testing the data. CA can function about the accuracy of the selected dataset. Precision is data accuracy that allows two events, namely 1 and 0. Recall serves to measure the ratio. F1 is the comparison of recall and precision. AUC is used to represent probabilities.

Table 2. Algorithm performance

Model	AUC	CA	F1	Precision	Recall
SVM	0.90	0.85	0.85	0.85	0.85
RF	0.91	0.84	0.84	0.84	0.84
NN	0.90	0.84	0.84	0.84	0.84

In Figure 8, the ROC analysis of each method is marked with a different color graph. RF are marked in Orange, SVMs in green, and NN in gray. In the three methods in one graph, it is clear that the sensitivity is high and specificity values are not more than 1 and when the target class value is changed, in all methods, the changes in sensitivity and specificity are not clearly visible.

From the analysis of the ROC curve in Figures 8 and 9, we can see the performance of the classification algorithm: the closer the curve is to the left and upper limits of the ROC space, the more accurate the classifier is. From the ROC analysis, it can be seen that the SVM model has a slightly better account of the classifier than the NN and RF models. This can be seen in eachclass, it can be seen that the curve of the average SVM approaches the axis or Y axis which indicates that SVM has an optimal classifier accuracy.

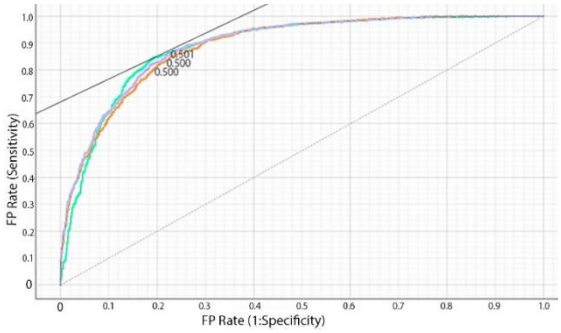


Figure 8. ROC analysis target class 0

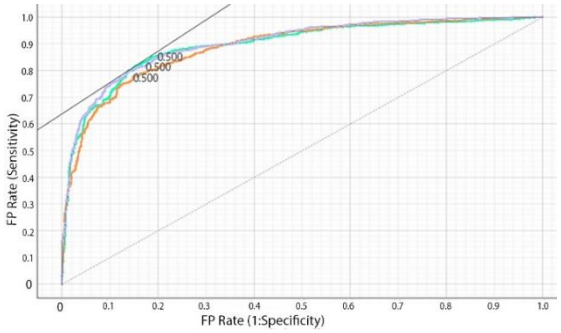


Figure 9. ROC analysis target class 1

4. CONCLUSION

This study evaluates the effectiveness of three device learning algorithms, RF, SVM, and NN in predicting coronary heart disease using a dataset of patients. Via applying information mining strategies with Orange programs and other exploration strategies, the dataset contains attributes with nominal and numeric facts, with the goal magnificence indicating the absence (0) or presence (1) of coronary heart disorder. The comparison consequences reveal that a few of the classification algorithms used SVM, random wooded area, and neural community, the SVM set of rules is the most correct and appropriate for predicting coronary heart sickness, accomplishing an accuracy charge of 85%. Random woodland set of rules excels at coping with complicated and sundry data, successfully lowering the danger of overfitting. Via combining outcomes from a couple of decision bushes, RF affords strong and correct predictions. SVM is particularly powerful in facts class by way of finding the most useful hyperplane that separates information into classes. Its ability to handle abnormal or overlapping statistics makes SVM a reliable tool in coronary heart disease analysis. NN have strong adaptive abilities to apprehend patterns and trends in large and complex datasets. With multiple layers of neurons, this algorithm can provide enormously correct predictions, although it calls for more time and computational sources. Integrating these 3 algorithms right into a heart sickness prediction gadget not handiest complements diagnostic accuracy but also aids in higher clinical choice-making. Using machine mastering in healthcare holds massive capacity for decreasing human errors, enhancing performance, and handing over better fitness effects for patients.

Even though SVM indicates the fine performance on this have a look at, there are several regions for in addition exploration. First, destiny studies could include evaluating newer algorithms or combining algorithms (ensemble methods) to look if accuracy can be improved. Second, including more attributes or variables to the dataset could enhance prediction nice. Subsequently, similarly studies have to do not forget the interpretability of models to ensure that prediction consequences can be understood and relied on by way of clinical practitioners in medical selection-making.

ACKNOWLEDGEMENTS

The authors would like to thank all parties STMIK Sinar Nusantara Surakarta. His research interests are in Decision Support System, Information Systems and Information Technology, University of Raharja, Tangerang, Bandung Technology University, Universitas Dian Nusantara Jakarta, the University of Al-Khairiyah Cilegon, Bina Insani universityi Bekasi, IBM Bekasi, Indonesia who have supported and funded this research.

REFERENCES

[1] A. Avdić, E. Kajan, D. Janković, and D. Avdić, "Towards Context-Aware Smart Healthcare Platform," *Ijeec - International Journal of Electrical Engineering and Computing*, vol. 3, no. 1, 2019, doi: 10.7251/ijeec1901026a.

[2] K. Angappan, N. Meenakshi, J. Evans E., H. Bharanika, and S. Jothi, "A Hybrid Approach for Heart Disease Prediction," *Proceedings of the First International Conference on Computing, Communication and Control System*, 2021, doi: 10.4108/eai.7-6-2021.2308784.

[3] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, D. S. Rajput, R. Kaluri, and G. Srivastava, "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis," *Evolutionary Intelligence*, vol. 13, pp. 185–196, 2020, doi: 10.1007/s12065-019-00327-1.

[4] R. P. Cherian, N. Thomas, and S. Venkitachalam, "Weight optimized neural network for heart disease prediction using hybrid lion plus particle swarm algorithm," *Journal of Biomedical Informatics*, vol. 110, 2020, doi: 10.1016/j.jbi.2020.103543.

[5] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 141, pp. 19–26, 2017, doi: 10.1016/j.cmpb.2017.01.004.

[6] S. Sharma and M. Parmar, "Heart Diseases Prediction using Deep Learning Neural Network Model," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 3, pp. 2244–2248, 2020, doi: 10.35940/ijitee.c9009.019320.

[7] K. A. Dauda, K. O. Olorede, and S. A. Aderoju, "A novel hybrid dimension reduction technique for efficient selection of bio-marker genes and prediction of heart failure status of patients," *Scientific African*, vol. 12, 2021, doi: 10.1016/j.sciaf.2021.e00778.
[8] V. V. Shankar, V. Kumar, U. Devagade, V. Karanth, and K. Rohitaksha, "Heart Disease Prediction Using CNN Algorithm," *SN Computer Science*, 2020, doi: 10.1007/s42979-020-0097-6.
[9] H. Henderi, R. Irawatia, I. Indra, D. A. Dewi, and T. B. Kurniawan, "Big Data Analysis using Elasticsearch and Kibana: A Rating Correlation to Sustainable Sales of Electronic Goods," *HighTech and Innovation Journal*, vol. 4, no. 3, pp. 583–591, 2023, doi: 10.28991/HIJ-2023-04-03-09.
[10] L. Yahaya, N. D. Oye, and E. J. Garba, "A Comprehensive Review on Heart Disease Prediction Using Data Mining and Machine Learning Techniques," *American Journal of Artificial Intelligence*, vol. 4, no. 1, 2020, doi: 10.11648/j.ajai.20200401.12.
[11] Y. Khourdifi and M. Bahaj, "The Hybrid Machine Learning Model Based on Random Forest Optimized by PSO and ACO for Predicting Heart Disease," in *ICCWCS 2019*, 2019, doi: 10.4108/eai.24-4-2019.2284088.
[12] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai and R. S. Suraj, "Heart Disease Prediction using Hybrid machine Learning Model," *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, 2021, pp. 1329-1333, doi: 10.1109/ICICT50816.2021.9358597.
[13] Mulyawan, A. Bahtiar, G. Dwilestari, F. M. Basysyar, and N. Suarna, "Data mining techniques with machine learning algorithm to predict patients of heart disease," *IOP Conference Series: Materials Science and Engineering*, 2021, doi: 10.1088/1757-899x/1088/1/012035.
[14] Y. Syahidin, A. P. Ismail, and F. N. Siraj, "Application of Artificial Neural Network Algorithms to Heart Disease Prediction Models with Python Programming," *Jurnal E-Komtek (Elektro-Komputer-Teknik)*, vol. 6, no. 2, 2022, doi: 10.37339/e-komtek.v6i2.932.
[15] H. Hamdani, H. R. Hatta, N. Puspitasari, A. Septiarini, and H. Henderi, "Dengue classification method using support vector machines and cross-validation techniques," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 3, pp. 1119-1129, 2022, doi: 10.11591/ijai.v11.i3.pp1119-1129.
[16] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning," *Computational Intelligence and Neuroscience*, 2021, doi: 10.1155/2021/8387680.
[17] N. M. Lutimath, C. Chethan, and B. S. Pol, "Prediction of heart disease using machine learning," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, pp. 474–477, 2019, doi: 10.35940/ijrte.B1081.0982S1019.
[18] K. Subhadra and B. Vikas, "Neural network based intelligent system for predicting heart disease," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, no. 5, pp. 484–487, 2019.
[19] T. L. Octaviani and d Z. Rustam, "Random forest for breast cancer prediction," in *AIP Conference Proceedings*, AIP Publishing, 2019, doi: 10.1063/1.5132477.
[20] D. C. Yadav and S. Pal, "Prediction of heart disease using feature selection and random forest ensemble method," *International Journal of Pharmaceutical Research*, vol. 12, no. 4, pp. 56–66, 2020, doi: 10.31838/ijpr/2020.12.04.013.
[21] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996, doi: 10.1007/BF00058655.
[22] R. Devika, S. V. Avilala, and V. Subramaniyaswamy, "Comparative study of classifier for chronic kidney disease prediction using naive bayes, KNN and random forest," *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2019, pp. 679-684, doi: 10.1109/ICCMC.2019.8819654.
[23] M. A. Jabbar and S. Samreen, "Heart disease prediction system based on hidden naïve bayes classifier," *2016 International Conference on Circuits, Controls, Communications and Computing, I4C 2016*, 2016, pp. 1-5, doi: 10.1109/CIMCA.2016.8053261.
[24] M. Akgül, Ö. E. Sönmez, and T. Özcan, "Diagnosis of heart disease using an intelligent method: A hybrid ANN – GA approach," *Advances in Intelligent Systems and Computing*, vol. 1029, pp. 1250–1257, 2020, doi: 10.1007/978-3-030-23756-1_147.
[25] R. R. Aswathi, K. P. Kumar, and B. Ramakrishnan, "Whale optimization approach for early heart disease prediction based on FCM using DCNN," *Review of Computer Engineering Research*, vol. 10, no. 4, 2023, doi: 10.18488/76.v10i4.3549.
[26] I. Yunianto, A. Kurniawan, and M. M. Mutoffar, "Comparison of Decision Tree, KNN and Naïve Bayes Methods in Predicting Student Late Graduation In the Informatics Engineering Department, Institute Business XYZ," *Adpebi International Journal of Multidisciplinary Sciences*, vol. 1, no. 1, pp. 374–383, 2022.
[27] K. Wankhede, B. Wukkadada, S. Rajesh, and S. Nair, "Machine Learning Techniques for Heart Disease Prediction," in *2023 Somaiya International Conference on Technology and Information Management (SICTIM)*, IEEE, 2023, pp. 28–33, doi: 10.1109/SICTIM56495.2023.10104919.
[28] D. Hemalatha and S. Poorani, "Machine learning techniques for heart disease prediction," *Journal of Cardiovascular Disease Research*, vol. 12, no. 1, pp. 93–96, 2021.
[29] N. Bhatla and K. Jyoti, "An analysis of heart disease prediction using different data mining techniques," *International Journal of Engineering Research & Tecnology (IJERT)*, vol. 1, no. 8, 2012.
[30] N. G. Nia, E. Kaplanoglu, and A. Nasab, "Evaluation of artificial intelligence techniques in disease diagnosis and prediction," *Discover Artificial Intelligence*, 2023. doi: 10.1007/s44163-023-00049-5.

BIOGRAPHIES OF AUTHORS

Didik Setiyadi received the Master degree (Master in Informatics Engineering) from the STTI Benarif Indonesia, Jakarta, Indonesia graduate in 2003. Currently a permanent lecturer in the information system study program at the STMIK Sinar Nusantara Surakarta. His research interests are in decision support system, information systems and information technology. He is as Steering Committee of Association Auditor in the Central Management of the Indonesian Information and Communication Technology (PASTIKINDO), and as Certification Division of Indonesian Computer, Electronics and Instrumentation Support Society (INDOCEISS). Currently he is also a BNSP assessor for LSP Digital Technology in the computer sector. He can be contacted at email: ddk.setiyadi20@gmail.com.

Henderi received the Doctor degree (Doctor in Computer Science) from the Universitas Gadjah Mada, Yogyakarta, Indonesia, with the Dissertation "Model of Performance Evaluation System for Study Programs Based on LED K-means Clustering". He is a Professor of Computer Science in the Informatics Division University of Raharja, Tangerang, Indonesia. In addition, He is serving as Vice Rector in Academic Affair and Head of the research group on Artificial Intelligence. His research interests are in decision support system, data mining, business intelligence that are in information systems. He is as Deputy Chairman of Association Auditor in the Central Management of the Indonesian Information and Communication Technology (PASTIKINDO), and as Central Board of Indonesian Computer, Electronics and Instrumentation Support Society (INDOCEISS). He can be contacted at email: henderi@raharja.info.

Anrie Suryaningrat received the Master of Informatics Engineering - Business Intelligence from Universitas Raharja, Tangerang, Indonesia, with the Thesis "Determining CPU Capacity on Three-Tier Architecture through Performance Testing using Apache JMeter and Elastic Stack". He is a Lecturer in Informatics Engineering at Universitas Dian Nusantara, Jakarta, Indonesia. In addition, He is serving as Head of the Computer Lab. at Universitas Dian Nusantara, Jakarta, Indonesia. His research interests are in cloud computing, network engineering, cybersecurity, machine learning, and data analyst. He is currently as organization committee in Indonesia Honeynet Project since 2019. He can be contacted at email: anrie.suryaningrat@dosen.undira.ac.id.

Rulin Swastika received the Master degree (Master in Informatics Engineering) from the STIMIK ERESHA Jakarta, Indonesia graduate in 2013. Currently a permanent lecturer in the Informatics Engineering study program at the University of Al-Khairiyah, Cilegon, Indonesia. Her research interests are in decision support system, information systems, and information technology. Currently she is also a BNSP assessor for LSP Tematik Nusantara in the computer sector. She can be contacted at email: swastikarulin@gmail.com.

Saludin received the B.Eng. degree in electronic engineering, graduated in 1989, Master's degree in informatics engineering, graduated in 2003, Ph.D. degree in economic, graduated in 2008, currently a permanent lecturer in the informatics engineering study program at Bina Insani universityi Bekasi-Indonesia, actively conducts research in the fields of computer science and artificial intelligence. He can be contacted at email: saludin@binainsani.ac.id.

Muhamad Malik Mutoffar Bachelor program majoring in Informatics Engineering, graduated in 2012, Master of Management graduating in 2015, currently pursuing his final thesis for the Master of Informatics Engineering major at Langlang Buana University, currently a permanent lecturer in the informatics engineering study program at the Universitas Teknologi Bandung, as Assessor for BNSP Computer Science Competency LSP BPPTIK KOMINFO, actively conducts research in the fields of management and computer science, several times receiving research grants, PDP grant, PKM grant and Technology Product Dissemination Program grant from DRPM DIKTI. He can be contacted at email: malik@utb-univ.ac.id.

Imam Yunianto Doctoral Student in the Department of Computer Science, Faculty of Mathematics and Natural Sciences, IPB University, Bogor, Indonesia. He has been teaching Computer Science at the Informatics Engineering Study Program, Muhammadiyah Business Institute, Bekasi, since 2020. Apart from that, he also serves as Chair of the Jakarta Special Region of the Association of Indonesian Researchers and Publishers (APPEPINDO) and administrator of the IPB University Postgraduate Student Forum in the field of Information Technology and is a reviewer of seven Sinta-accredited journals. His research interests are in artificial intelligence, data mining and machine learning. He can be contacted at email: imamyunianto@gmail.com, imam@ibm.ac.id, and imam_s3ipb_komimam@apps.ipb.ac.id.